Subspace Techniques for Radio-Astronomical Data Enhancement

Sarod Yatawatta[1]
Kapteyn Institute, University of Groningen,
and ASTRON,
The Netherlands.
yatawatta@astron.nl

Abstract

Radio astronomical observations have very poor signal to noise ratios, unlike in other disciplines. On the other hand, it is possible to observe the object of interest for long time intervals as well as using a wider bandwidth. Traditionally, by averaging in time and in frequency, it has been possible to improve the signal to noise ratio of astronomical observations to improve the dynamic range. This is possible due to the inherent assumption that the object of interest in the sky is invariant over time and the frequency range of observation. However, in reality this assumption does not hold, due to intrinsic variation of the sky as well as due to errors generated by the instrument. In this paper, we shall discuss an alternative to averaging of images, without ignoring subtle changes in the observed data over time and frequency, using subspace decomposition. By separation of data to signal and noise subspaces, not only would this improve the quality of the data, but also enable us to detect faint artifacts due to calibration errors, interference etc.

I. Introduction

Astronomical observations have very weak signal to noise ratios (SNR). In radio astronomy (interferometry), this weak SNR is improved mainly by prolonging the observation time. This not only improves the SNR, but also sampling of the uv plane. A typical observation has a bandwidth of a few MHz, divided into smaller narrowband channels. Normally the bandwidth of a single channel is about a few kHz. After interference mitigation, calibration and imaging (also deconvolution) each of these narrow band channels, it is normal procedure to average them, on the image plane. The resultant averaged image has an improved SNR compared to images made by each of the narrowband channels, assuming all systematic errors have been removed. This is the typical process by which current radio telescopes like the Westerbork Synthesis Radio Telescope (WSRT), is able to achieve dynamic ranges in the order of 1,000,000 to 1.

However, there is a significant drawback in this averaging procedure, because it is based on a erroneous assumption: that the source as well as the effects of the instrument remains invariant in all averaged data. Only the noise varies and averaging significantly decreases the noise power. The invariance assumption is accurate to some extent, but closer scrutiny reveals that there is significant amount of variation:

- There is intrinsic variation of the sky (or the sources) with frequency (non zero spectral indices).
- The uv points scale with frequency, thus the point spread function (PSF) change in the images.
- Transient sources as well as radio frequency interference (RFI) create temporal variation.
- Calibration errors, small as they might be, create subtle variations in time as well as in frequency.

Indeed, by averaging, both noise and the aforementioned effects get suppressed. However, it is to our advantage that we detect such variations in our images. For instance, we could detect trace amounts of RFI and calibration errors etc., that appear only in certain time or frequency ranges. This could also enable us to discard certain parts of the data, where such effects are dominant, and thus improve the quality of our averaged, final image.

Subspace techniques have previously being used in image compression [1], [2], as well as image comparison. However, to the best of our knowledge, this has not been used in (radio) astronomical image processing.

The rest of the paper is organized as follows: In the next section, we give a formal representation of images and subspace decomposition. Next, we investigate various uses of the image subspace decomposition. Later, we consider possible extensions of the proposed method, by taking into account the data weights as well as the affect of PSF. Finally, we give our conclusions.

Notation: All bold lowercase letters represent a column vector, all bold uppercase letters represent a matrix.

[1]This work was first presented in the SKA calibration and imaging workshop, Cape Town, SA, December 2006 and in URSI General Assembly, Chicago, US, August 2008

II. Mathematical Preliminaries

Let us represent an arbitrary image, i.e., a set of pixels with arbitrary arrangement in two dimensions, as a vector $\mathbf{b}$. Let the number of chosen pixels be N, so the dimension of vector $\mathbf{b}$ is $N \times 1$. Instead of one such image, we have a set of images, of the same part of the sky. Let us denote this set by $\mathcal{I}$. Let us assume we have M such images, which could differ in terms of any of the following:

- Observing frequency, one image per each channel.
- Observing time, different observation dates.
- Observing instrument, using different radio telescopes, i.e. WSRT or the Very Large Array (VLA).
- Integration time, using snapshots

In other words, the images in the set $\mathcal{I}$ have one or more forms of diversity. Let $\mathbf{b}_j$ represent the j-th such image from $\mathcal{I}$. We assume all the images in $\mathcal{I}$ could be represented by an arbitrary orthonormal basis $\mathbf{V}$ ($N \times N$ matrix), as in (1). Note that $\mathbf{V}$ is used to represent the components visible in the image, i.e., the signal, not the noise.

$$\mathbf{b}_j = \mathbf{V}\mathbf{a}_j + \mathbf{n}, \quad j \in [1 \dots M] \tag{1}$$

In (1), $\mathbf{a}_j$ represent the components of the basis $\mathbf{V}$ used to represent the j-th image. The noise, which is assumed to be independent in each image in $\mathcal{I}$ is represented by $\mathbf{n}$ ($N \times 1$ vector). Moreover, the signal or the artifacts are represented by $\mathbf{V}\mathbf{a}_j$ in (1).

We make the following assumptions:

- The variations in all the images in $\mathcal{I}$ is small enough to be represented by at most N orthonormal vectors.
- The noise in each image a stationary random process with autocorrelation $\sigma^2\mathbf{I}$, and is uncorrelated between pixels as well as with the signal.
- The noise power is much lower compared to the signal or artifacts.

Let us consider the autocorrelation $\mathbf{R}$ of the images, given by (2).

$$\mathbf{R} \stackrel{\triangle}{=} E\{\mathbf{b}\mathbf{b}^T\} \tag{2}$$

Using M images, the estimate of the autocorrelation $\hat{\mathbf{R}}$ is given by (3).

$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{j=1}^{M} \mathbf{b}_j \mathbf{b}_j^T \tag{3}$$

We exploit the assumption that noise is uncorrelated with signal, i.e., $E\{\mathbf{n}\mathbf{a}^T\} = E\{\mathbf{a}\mathbf{n}^T\} = \mathbf{0}$. Because $\mathbf{V}\mathbf{V}^T = \mathbf{I}$, and we get

$$\mathbf{R} = \mathbf{V}E\{\mathbf{a}\mathbf{a}^T\}\mathbf{V}^T + \sigma^2\mathbf{I} = \mathbf{V}\left(E\{\mathbf{a}\mathbf{a}^T\} + \sigma^2\mathbf{I}\right)\mathbf{V}^T. \tag{4}$$

From (4), we see that the column space of $\mathbf{R}$ is a subset of the column space $\mathbf{V}$.

$$col(\mathbf{R}) \subseteq col(\mathbf{V}) \tag{5}$$

Using the Singular Value Decomposition (SVD) we can find an equivalent column space $\mathbf{R} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T$ as in (6).

$$col(\mathbf{U}) \subseteq col(\mathbf{V}) \tag{6}$$

The singular values in $\boldsymbol{\Sigma}$ will characterize the signal and noise subspace eigenmodes. Because the noise has much lower power compared with the signal, the eigenmodes corresponding to the dominant singular values in $\boldsymbol{\Sigma}$ will represent the subspace of $\mathbf{V}$.

Hence, we can divide eigenvectors as noise or signal according to their singular values (7), where $\mathbf{U}_s$ and $\mathbf{U}_n$ represent the signal and noise subspaces, respectively.

$$\mathbf{U} = \left[\begin{array}{ccc} \mathbf{U}_s & \vdots & \mathbf{U}_n \end{array}\right] \tag{7}$$

Given an image $\mathbf{b}_j$, we can separate the signal $\mathbf{b}_{sj}$ and noise $\mathbf{b}_{nj}$ as in (8).

$$\mathbf{b}_{sj} = \mathbf{U}_s\mathbf{U}_s^T\mathbf{b}_j \quad \mathbf{b}_{nj} = (\mathbf{I} - \mathbf{U}_s\mathbf{U}_s^T)\mathbf{b}_j \tag{8}$$

Note in practice, in (3), we normally have fewer images than the image size in pixels, i.e., $N \gg M$. Hence, we would only have M non zero singular values in $\boldsymbol{\Sigma}$. However, by our assumption (small enough variation), it should be possible to represent the signal space using fewer eigenmodes than M.

III. Applications

The aforementioned technique has been widely used in image compression and comparison in other disciplines. However, we could find novel uses of this technique in radio astronomy.

- Enhancement of signal to noise ratio: For each image $\mathbf{b}_j$ in the set $\mathcal{I}$, we get the enhanced version $\mathbf{b}_{sj}$ in (8). These images could be used for further processing, like estimation of spectral indices of fainter sources.
- Automatic image classification: A typical observation (say in LOFAR) consists of a few thousand channels, with an image being made for each one. It is assumed the majority of these images are good quality, with proper calibration and imaging. However, a few of these images would still have faint RFI or calibration errors. In order to automatically determine image quality, we could use the power of the noise component, i.e., the norm of $\mathbf{b}_{nj}$ in (8). For a poor quality image, this should be much higher compared to a good quality image.
- Continuum subtraction, spectral lines: Let us assume that the j-th image contains a weak spectral line, hidden under a strong continuum signal that is common to all the images in the set $\mathcal{I}$. Just by looking at this image $\mathbf{b}_j$, it might not be possible to observe this spectral line because of the continuum signal. However, by subtracting this, in $\mathbf{b}_{nj}$, we should be able to observe this.
- Detection of faint, narrowband RFI: The same technique applied to spectral lines could be applied to detect faint, narrowband RFI. It is assumed that the RFI affects only a few images in the set $\mathcal{I}$.

An implementation of this method which is capable of decomposing FITS files can be downloaded from [3].

IV. Example

In this example, we have considered a WSRT observation with faint narrowband RFI, which has not been detected by the flagging algorithm. We have about 32 channels of data, each with a bandwidth of 100 kHz centered around 140 MHz. We have applied the proposed technique on this dataset. The results are given on Fig. 1.

In order to reduce the computational cost (for an image with N pixels, the cost of computation of the SVD is $O(N^3)$), we have used a divide and conquer approach. Given a set of images with a large number of pixels, we first divide them into smaller sub-images, using an arbitrary boundary. For each set of these sub images (one taken from each large image), we apply the proposed technique. This saves the computational cost significantly. However, this could also lead to new artifacts. For instance, the tile like structure appearing on Fig. 1 (d) is due to these sub-images. But as seen from this image, these artifacts are less significant to affect the outcome.

V. Conclusions

We have proposed a technique to improve the quality of the data in astronomical observations using subspace decomposition. Application of this technique to real data has shown us that it is superior to the normally used averaging techniques.

VI. Acknowledgement

The author would like to thank Ger de Bruyn and Jan Noordam for posing this problem. This work was supported by LOFAR and SNN. LOFAR is being funded by the European Union, European Regional Development Fund, and by “Samenwerkingsverband Noord-Nederland”, EZ/KOMPAS.

References

[1] H. C. Andrews and C. L. Patterson, “Singular value decompositions and digital image processing,” *IEEE Trans. ASSP*, no. 24, pp. 26-53, 1976.

[2] J. F. Yang and C. L. Lu, “Combined techniques of singular value decomposition and vector quantization for image coding,” *IEEE Trans. on Image Processing*, no. 4, pp. 1141-1146, 1995.

[3] \protect\vrulewidth0pthttp://www.astro.rug.nl∼yatawatta/software/

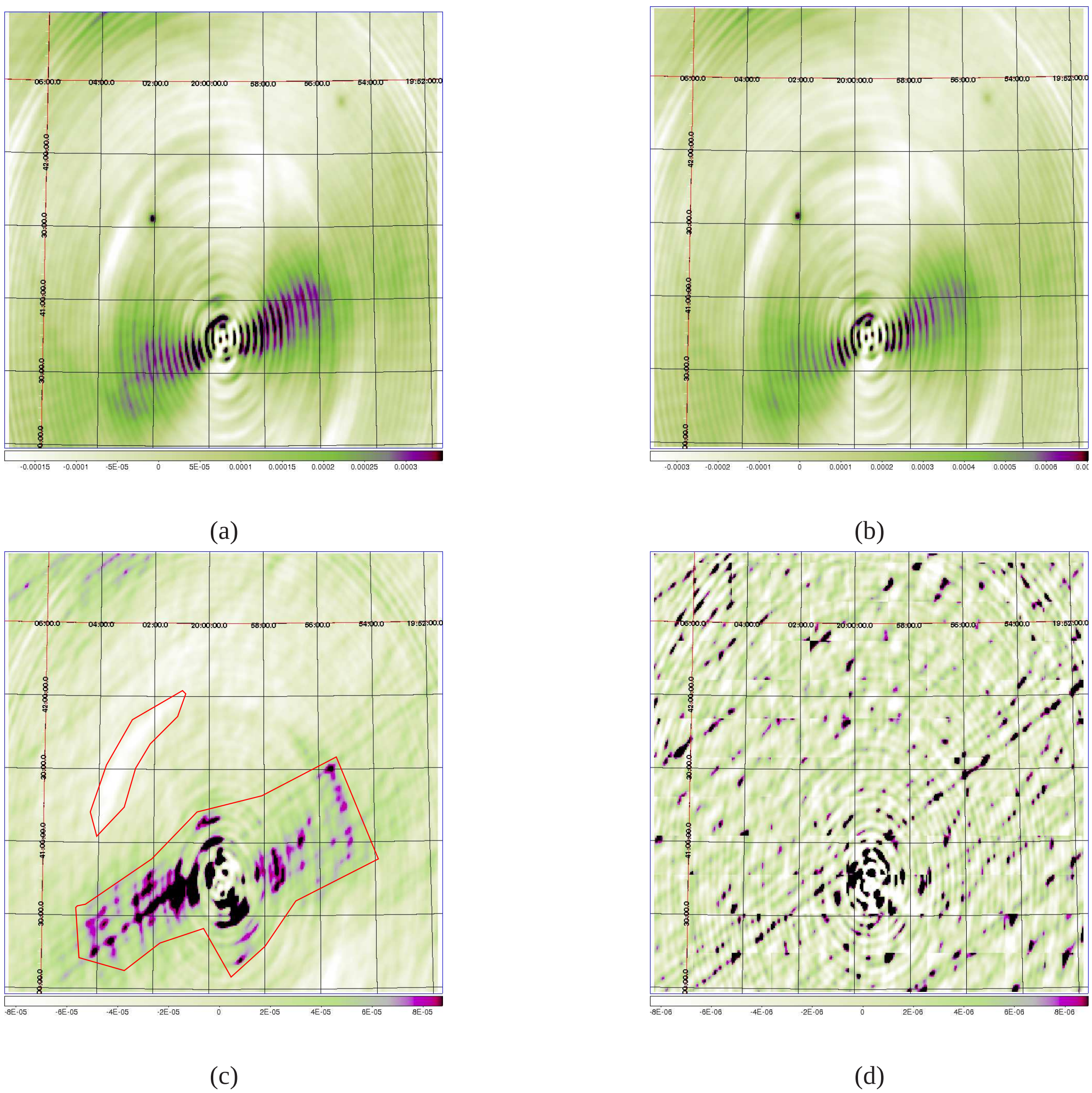

Fig. 1. Partially deconvolved residual image around Cyg A using WSRT data: {(a)} residual image made by a single channel of data, {(b)} mean image using 32 channels, {(c)} difference image between (a) and (b), {(d)} residual remaining after removal of signal subspace from the image {(a)}. It is clearly evident that the image (a) contains faint RFI, seen on (d) as diagonal lines in black. These appear also on the difference image (c) but it is hard to distinguish them because of the artifacts remaining. The artifacts on (c) are outlined using the red polygons. By using the proposed technique, we are able to remove most of these artifacts. The tiled structure on (d) is due to the divide-and-conquer approach in applying this technique.